

University of Groningen

Realisme en convergentie, of hoe het succes van de natuurwetenschappen verklaard moet worden

Kuipers, Theo A.F.

Published in:
EPRINTS-BOOK-TITLE

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
1991

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Kuipers, T. A. F. (1991). Realisme en convergentie, of hoe het succes van de natuurwetenschappen verklaard moet worden. In *EPRINTS-BOOK-TITLE* (blz. 61-83). Koninklijke Van Gorcum.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

4 Realisme en convergentie of hoe het succes van de natuurwetenschappen verklaard kan worden

Theo A.F. Kuipers

Deze bijdrage bestaat uit drie delen. In het eerste deel wordt een overzicht gegeven van een aantal pogingen om een formele uitleg te geven van de idee dat theorieën kunnen convergeren naar de waarheid. In het tweede deel worden de hoofdlijnen van één van deze explicaties geschetst. Deze sluit enerzijds nauw aan bij de ideeën van Popper, maar houdt anderzijds enkele fundamentele koerswijzigingen in. In het derde deel worden enkele belangrijke formele aspecten van deze explicatie gepresenteerd. In dit laatste deel en in de appendices is kennis van de elementaire verzamelingenleer voorondersteld.

1 Theorieën over waarheidsgelijkenis en convergentie

1.1 Waarheidsgelijkenis

Toen Popper (1962:233) zijn beroemde definitie van 'dichter bij de waarheid' gaf was dat een weinig schokkende intellectuele gebeurtenis. Iedereen kon reageren met te stellen dat die definitie was zoals hij zou moeten zijn en zelfs dat geen andere verwacht had kunnen worden. De definitie was immers uiterst plausibel: een theorie is dichter bij de waarheid dan een andere als de eerste enkele ware consequenties meer, en enkele onware consequenties minder, heeft dan de tweede.

Ongeveer tien jaar later werd het met terugwerkende kracht alsnog een schokkende gebeurtenis, toen Miller en Tichý onafhankelijk van elkaar bewezen dat een onware theorie, dat is een theorie met tenminste één onware consequentie, volgens Poppers definitie nooit dichter bij de waarheid kon zijn dan een andere. Voor een bewijs, zie *Appendix 1*.¹

Dit bewijs liet zien dat Poppers definitie irrelevant was, omdat zij geen recht kon doen aan een basisintuïtie over de hoogtepunten in de wetenschapsgeschiedenis, namelijk dat nieuwe theorieën, zoals die van Einstein, succesvoller zijn in het verklaren van empirische feiten en wetten dan hun voorgangers, zoals die van Newton, juist omdat ze, hoewel vermoedelijk eveneens onwaar, niettemin dichter bij de waarheid zijn. Anders gezegd, het

1. Zie de 'Introduction' in Kuipers (1987a) voor de vindplaatsen van de oorspronkelijke en enkele andere bewijzen van deze stelling. Het bewijs in Appendix 1 is daar een variatie op.

grotere succes van nieuwe theorieën moet niet zozeer verklaard worden in termen van de vraag of de theorieën waar of onwaar zijn, maar in termen van de afstand van die theorieën tot de waarheid: de vermoedelijk onware theorieën convergeren naar de waarheid. Onder 'de waarheid' wordt daarbij verstaan de onbekende ware theorie over het betreffende domein van verschijnselen.

Miller en Tichý ontketenden met hun stelling, die immers een convergente serie onware theorieën uitsluit, naast teleurgestelde gezichten vooral sceptische reacties, zoals 'alleen het intuïtieve idee is belangrijk', 'zie je wel dat je geen filosofische problemen met formele middelen kunt oplossen' en, *last but not least*, 'het is de straf voor het vrijelijk spreken over *de waarheid*'.

Na enige tijd echter bekwamen Miller, Tichý en andere filosofen van de schrik en begonnen alternatieve definities voor 'dichter bij de waarheid' ('verisimilitude') te ontwikkelen. Thans zijn er in de internationale literatuur tenminste vier formele benaderingen te onderscheiden van 'verisimilitude' of 'waarheidsgelijkenis' (*truthlikeness*) zoals het onderwerp tegenwoordig meestal wordt genoemd. Deze benaming verwijst naar het feit dat de afstand van een theorie tot de ware theorie gaandeweg is geherinterpreteerd als de mate waarin die theorie met de ware theorie gelijkenis vertoont. Uitleggen wat er precies onder 'gelijkenis tussen theorieën' moet worden verstaan is de kern van het probleem. Het begrip 'toenemende of convergerende gelijkenis' volgt dan meestal vanzelf.

De vier recente explicaties van waarheidsgelijkenis zijn alle in staat de klip te omzeilen waarop Poppers definitie schipbreuk leed, zodat een serie onware theorieën die convergeert naar de waarheid technisch mogelijk is.

Vanwege deze eigenschap kan met alle vier explicaties de claim geformuleerd worden dat historisch opeenvolgende theorieën doorgaans een serie naar de waarheid convergerende theorieën vormen. Dat wil echter nog niet zeggen dat ze alle vier ook een duidelijke onderbouwing kunnen geven van de eerder genoemde basisintuïtie dat toenemend succes van theorieën verklaard moet kunnen worden in termen van toenemende waarheidsgelijkenis. Het onderbouwen van die intuïtie zou de uitdaging van Laudan (1981) kunnen worden genoemd. Volgens Laudan gaan convergentierealistenaars altijd uit van dat verband, maar ze zouden het moeten bewijzen. Slechts twee van de vier genoemde explicaties pretenderen aan die uitdaging te voldoen. Niiniluoto's antwoord, dat een kwantitatief waarheidsgelijkenis-begrip vooronderstelt, blijkt alleen op het eerste gezicht plausibel en er rust bovendien een flinke hypotheek op. Mijn kwalitatieve explicatie, die in deel 2 en 3 behandeld wordt, leidt tot een antwoord aan Laudan dat zonder hypotheek is en

goed aansluit bij andere inzichten.

1.2 Convergentierealisme

Voordat ik overga op de afzonderlijke explicaties wil ik nog het gemeenschappelijke 'realistische' uitgangspunt van alle vier explicaties (en ook van Poppers mislukte poging) kort aanduiden. Dit gemeenschappelijk realisme betreft een gematigde vorm van 'metafysisch realisme': 'de waarheid' over (een deel van) de werkelijkheid wordt geacht enerzijds bepaald te worden door de aard van de onafhankelijke werkelijkheid zelf, maar anderzijds ook door het gekozen conceptueel raamwerk (de gekozen taal) en wel in die zin dat het gekozen conceptueel raamwerk bepaalt welke aspecten van de werkelijkheid tot uitdrukking kunnen komen. Als vanzelfsprekend hoort bij deze benadering dat er meerdere conceptuele raamwerken mogelijk zijn en dus meerdere 'waarheden'. Thermodynamica en statische mechanica bijvoorbeeld zijn vanuit twee verschillende conceptualiseringën op zoek naar 'hun eigen waarheden' (waarbij overigens ook sprake is van interessante interacties). Extreem metafysisch realisme daarentegen gaat er vanuit dat de werkelijkheid haar (ideale) conceptualisering met zich meebrengt, en dat er dus eigenlijk slechts sprake kan zijn van één waarheid.² Wanneer gematigd realisme wordt aangevuld met het idee van 'naar de waarheid convergerende theorieën' wordt gesproken van convergentierealisme.

Gematigd realisme hoeft niet te ontaarden in (conceptueel) relativisme zolang maar aangenomen wordt dat er allerlei verbanden tussen conceptuele raamwerken kunnen bestaan en dat zulke raamwerken kunnen worden samengevoegd tot meer omvattende raamwerken. Kortom, relativisme kan worden vermeden door de zogenaamde incommensurabiliteit tussen paradigma's, onderzoeksprogramma's, raamwerken en dergelijke op formele gronden af te wijzen.³

Gematigd realisme is goed verdedigbaar voor de natuurwetenschappen: het impliceert uiteraard dat theorieën waar of onwaar zijn en dat theoretische termen pretenderen te verwijzen naar iets in de werkelijkheid. Het is echter problematisch voor de sociale wetenschappen, omdat niet alleen de

2. Tegenover 'metafysisch realisme' staat 'intern realisme' waarbij de werkelijkheid zogenaamd epistemologisch wordt geconstitueerd.

3. Technisch kan dit op grond van het triviale inzicht dat het altijd mogelijk is twee termen die een bijna maar niet geheel gelijke betekenis hebben niet te identificeren, maar als twee verschillende termen te handhaven in een overkoepelend raamwerk. Zo'n raamwerk kan door eenvoudige samenvoeging verkregen worden, waarbij alleen twee termen met volstrekt gelijke betekenis worden geïdentificeerd.

conceptuele representatie van de sociale werkelijkheid, maar ook de sociale werkelijkheid zelf een menselijke constructie is. Met andere woorden, terwijl aangenomen mag worden dat er een onafhankelijke natuurwerkelijkheid bestaat, lijkt het niet zinvol om over een onafhankelijke sociale werkelijkheid te spreken.

1.3 Twee syntactische benaderingen

De benadering van Schurz en Weingartner (1987) is de meest letterlijke poging tot reparatie van Poppers oorspronkelijke idee. Dat luidde in gelijkentermen geformuleerd: een theorie lijkt meer op de ware theorie dan een andere als de eerste meer ware en minder onware consequenties heeft dan de tweede. Door de verzameling consequenties van een theorie in te perken tot de zogenaamde 'gereduceerde relevante (RR-)consequenties' blijkt het wel mogelijk dat een onware theorie meer lijkt op de waarheid dan een andere. Om enig idee te geven van een RR-consequentie, stel bijvoorbeeld dat bewering β logisch volgt uit α , dan heet $\beta \vee \gamma$, een disjunctie die nu voor willekeurige γ volgt uit α , een irrelevante consequentie van α . Hoewel het (syntactische) begrip van een RR-consequentie, waarvan hier de details niet verder ter zake doen, voor andere doeleinden is ontwikkeld en dus geen *ad hoc* constructie voor het onderhavige probleem is, heeft de toepassing hier toch een erg *ad hoc* karakter: het voldoet aan de technische eis, zonder verhelderende meerwaarde voor het onderwerp waarheidsgelijkenis.

Tichý en Oddie (Oddie 1986, 1987) gaan uit van een idiosyncratische versie van de logisch-modeltheoretische opvatting van theorieën waarin theorieën worden opgevat als verzamelingen modellen van beweringen. In hun idee over waarheidsgelijkenis staat echter de syntactische gelijkheid tussen de complexe beweringen die de theorieën weergeven centraal. Dit leidt volgens hen tot een soort afbeeldingstheorie *à la* Wittgenstein. Hoofdprobleem van deze explicatie is dat zij technisch zeer ingewikkeld is, zonder dat duidelijk gemaakt wordt dat dat echt nodig is voor het onderhavige probleem. Het specifieke technische karakter lijkt zelfs implausibel omdat onderzoekers zich bij het vergelijken van natuurwetenschappelijke theorieën absoluut niet plegen te verdiepen in de syntactische structuur van de betreffende beweringen.

1.4 De benadering van Niiniluoto

De twee bovengenoemde explicaties hebben gemeen dat ze primair syntactisch zijn en dat ze geen plausibel antwoord op Laudans uitdaging hebben.

De explicatie van Niiniluoto (1987a,1987b) is niet per definitie syntactisch en leidt wel tot een antwoord aan Laudan. Niiniluoto stelt de abstracte vorm van een cognitief probleem centraal, waarbij de hamvraag is wat de correcte oplossing van het probleem is. Een voorbeeld van een cognitief probleem is uiteraard de vraag, binnen een zekere context, welke theorie de ware theorie is. Niiniluoto gaat uit van een gegeven verzameling potentiële antwoorden die elkaar onderling uitsluiten en die tezamen alle mogelijkheden uitputten, de zogenaamde complete potentiële (CP-)antwoorden. Precies één van die CP-antwoorden is het correcte antwoord. De meta-vraag is dan of het ene CP-antwoord dichterbij het correcte CP-antwoord is dan het andere. Deze meta-vraag kan eenvoudig worden uitgebreid tot incomplete potentiële antwoorden, dat zijn potentiële antwoorden die bestaan uit een disjunctie van CP-antwoorden. In combinatie met een voor het probleem geschikte afstandsfunctie leidt dit tot een kwantitatieve definitie van de waarheidsgelijkenis (de gelijkenis met het CP-antwoord) van ieder potentieel antwoord.

Als de antwoorden expliciet in termen van beweringen, dus syntactisch geformuleerd zijn leidt de aanpak van Niiniluoto tot de syntactische definitie van waarheidsgelijkenis voor het betreffende cognitieve probleem. Maar als de antwoorden anderssoortig geformuleerd zijn, bijvoorbeeld in termen van (intervallen van) getallen - om een favoriet voorbeeld van Niiniluoto te noemen - ontstaat ook een anderssoortige definitie. Zolang er een hanteerbare syntactische formulering mogelijk is gaat daar Niiniluoto's voorkeur wel naar uit.

Een bedenking tegen Niiniluoto's aanpak is dat hij meteen wil doorstoten naar een kwantitatief begrip van waarheidsgelijkenis. De prijs daarvoor is de noodzaak een beroep te doen op kwantitatieve afstandsfuncties die in de onderhavige context vrijwel altijd een zeer arbitrair karakter hebben.

De kwantitatieve aanpak is voor Niiniluoto echter noodzakelijk om zijn antwoord aan de uitdaging van Laudan te kunnen formuleren. Dit antwoord klinkt op het eerste gezicht voor het overige aantrekkelijk, als men tenminste een of andere vorm van niet-objectieve waarschijnlijkheid, zoals subjectieve of inductieve waarschijnlijkheid, niet bij voorbaat afwijst. De basisidee is dat de waarheidsgelijkenis van een theorie in een reële wetenschappelijke context weliswaar niet kan worden vastgesteld, omdat het correcte antwoord onbekend is, maar dat het wel mogelijk is een schatting van de waarheidsgelijkenis te maken op grond van de beschikbare empirische data. Die schatting wordt technisch uiteraard gedefinieerd als de verwachtingswaarde van de waarheidsgelijkenis.

De onderbouwing van de basisintuïtie verloopt nu als volgt: als een

theorie een grotere waarheidsgelijkenis heeft dan een andere dan zullen de empirische data in de regel ook zo zijn dat de op grond van die data geschatte waarheidsgelijkenis van de eerste theorie groter is dan die van de tweede.

Op deze onderbouwing rust een zware hypotheek en de plausibiliteit lijkt beperkt tot de voorgestelde explicatie van de notie 'geschatte waarheidsgelijkenis'. De hypotheek is dat voor ieder cognitief probleem in principe een waarschijnlijkheidsfunctie gedefinieerd moet kunnen worden die niet alleen geschikt is om de geponeerde verbanden te leggen, maar die bovendien zo in elkaar zit dat een theorie met een grotere geschatte waarheidsgelijkenis in de regel ook succesvoller is, want uiteindelijk ging het om het verklaren van succestoename. Maar Niiniluoto geeft geen duidelijke explicatie van de begrippen 'succes' en 'succesvoller'. Afgezien van de vraag of hierbij geen circulariteit zou ontstaan, geldt in elk geval dat in Niiniluoto's onderbouwing zijn overigens plausibele explicatie van de notie 'geschatte waarheidsgelijkenis' centraal staat. Dit is vreemd omdat Laudans uitdaging in beginsel alleen verlangt te laten zien dat grotere waarheidsgelijkenis meer succes impliceert. Het begrip 'grotere geschatte waarheidsgelijkenis' lijkt een verwarrende omweg, ook in de ogen van Laudan (1981:n.10).

2 Structuralistisch convergentierealisme: hoofdlijnen

De explicatie die ikzelf in eerdere publicaties heb uiteengezet sluit enerzijds aan bij de algemene probleem- en doelstelling van het convergentierealisme, maar houdt in een aantal opzichten koerswijzigingen in ten opzichte van Poppers oorspronkelijke opzet - die technisch schipbreuk leed - en de hiervoor geschetste voorstellen tot reparatie.

Evenals Popper en in tegenstelling tot met name Niiniluoto lijkt het mij verstandig eerst een kwalitatieve definitie van waarheidsgelijkenis vast te stellen. Op die manier is het mogelijk dat men zich eerst concentreert op ondubbelzinnige gevallen, zonder vertroebeling met allerlei arbitraire keuzes die onvermijdelijk zijn bij een kwantitatieve definitie.

2.1 Structuralistische aanpak

Omdat theorieën traditioneel worden opgevat als verzamelingen beweringen lijkt het voor de hand te liggen naar een definitie van waarheidsgelijkenis te zoeken in termen van beweringen, maar vanzelfsprekend is dat niet. Naast de zogenaamde *statement view* is er door Suppes en Sneed de zogenaamde

structuralistische benadering van theorieën ontwikkeld. In deze benadering wordt een theorie primair geïdentificeerd met de verzameling structuren⁴ die direct of indirect altijd met een theorie geassocieerd is.

Het gaat hier niet om twee onverenigbare opvattingen over de reconstructie van theorieën, zoals nogal eens wordt gesuggereerd. Het gaat juist om reconstructies die in beginsel verenigbaar en vergelijkbaar zijn. Voor het geval de beweringen van een theorie deel uitmaken van een zogenaamde eerste orde taal komt het er zelfs eenvoudig op neer dat de structuren van die theorie gevormd worden door de zogenaamde modellen van die theorie, dat zijn die structuren waarop precies alle beweringen van de theorie waar zijn.

Omdat theorieën dus kunnen worden opgevat als verzamelingen beweringen én als verzamelingen structuren is het enige dat telt de vraag welke reconstructie voor een bepaald probleem de meest bruikbare is. De claim dat de structuralistische opzet het meest geschikt is voor de problematiek van waarheidsgelijkenis is dus alleen te beoordelen op grond van de feitelijke behandeling van die problematiek vanuit dat perspectief: *the proof of the pudding is in the eating*.

De eerste aanzet in deze richting is gegeven door Miller (1978). Hij ontwikkelde een (kwalitatieve) definitie van waarheidsgelijkenis in termen van de modellen van een theorie, aangenomen dat die theorie eerste orde formuleerbaar is. De hier te presenteren explicatie, die in zuiver technische zin overeenstemt met die van Miller maar die onafhankelijk van Miller werd opgesteld, is niet alleen consequent structuralistisch maar er wordt ook fundamenteel gebroken met de zogenaamde *actual world doctrine*.

2.2 Descriptieve en theoretische waarheid

In de hele discussie over waarheidsgelijkenis is altijd als vanzelfsprekend aangenomen dat 'de waarheid' waarnaar theorieën geacht worden te convergeren betrekking heeft op 'de ware beschrijving van (het bestudeerde deel van) de actuele wereld'. Dat is vreemd omdat in andere contexten, bijvoorbeeld als het om het verklaren van individuele gebeurtenissen gaat, altijd wordt aangenomen dat theorieën dat alleen kunnen als er zogenaamde 'initiele condities' aan worden toegevoegd. Dit impliceert immers dat een theorie altijd onvoldoende specifieke informatie over de actuele wereld bevat om die

4. Een structuur bestaat uit één of meer verzamelingen met daarop gedefinieerde eigenschappen, relaties en functies.

wereld zonder meer te kunnen verklaren. In dit licht is het vreemd om er van uit te gaan dat theorieën wel naar het geheel van de specifieke informatie over, *in casu* de ware beschrijving van, de actuele wereld zouden convergeren.

Er zijn dan ook twee soorten waarheid en dus twee soorten waarheidsgelijkenis in het spel. 'De descriptieve waarheid' betreft de ware beschrijving van de actuele wereld die hoort bij de gehanteerde taal en dit roept de vraag op wanneer de ene beschrijving dichter bij de ware beschrijving is dan een andere. In structuralistische termen gaat het om de structuur die, gegeven de conceptuele middelen, de correcte representatie van de actuele wereld biedt en de vraag wanneer de ene structuur meer lijkt op de correcte structuur dan een andere.

Bij 'de theoretische waarheid' gaat het, gegeven de conceptuele middelen (een taal of anderszins) om de juiste karakterisering van de werelden die empirisch mogelijk zijn, *in casu* de ware theorie en is de vraag of de ene theorie meer lijkt op de ware theorie dan andere.

In de waarheidsgelijkenisdiscussie zijn de problemen van descriptieve en theoretische waarheidsgelijkenis vaak verward. Beide problemen kunnen overigens heel goed worden opgevat als (kwalitatieve varianten van) cognitieve problemen in de zin van Niiniluoto. In eerste instantie (Kuipers:1982) meende ik dat het probleem van descriptieve waarheidsgelijkenis het beste in termen van beweringen kon worden opgelost, bij nader inzien (Kuipers: 1987d) kan de aanzet van toen ook opgevat worden als een aanzet voor een structuralistische oplossing van dit probleem.

Hoe dit ook zij, het probleem van theoretische waarheidsgelijkenis is veel interessanter omdat het in de natuurwetenschappen primair gaat om vast te stellen wat empirisch mogelijk is en wat empirisch onmogelijk. Theorieën zijn daarop gericht. Zoals we nog zullen zien wordt ook het succes van theorieën *de facto* in die termen gemeten. Ik noem deze opvatting 'theoretisch realisme'.

Vanuit het structuralistisch perspectief komt het probleem van theoretische waarheidsgelijkenis neer op het vergelijken van verzamelingen van conceptuele mogelijkheden, dat zijn (verzamelingen van) structuren die door de conceptuele middelen karakteriseerbaar zijn. Met iedere theorie is een verzameling conceptuele mogelijkheden geassocieerd; voor de ware theorie is dat de verzameling empirische mogelijkheden, aangeduid met X. Theorie en geassocieerde verzameling duiden we gemakshalve met dezelfde (hoofd)letter aan. Het symmetrisch verschil tussen twee verzamelingen is algemeen gedefinieerd als de verzameling van die elementen die slechts tot één van de twee

verzamelingen behoren, waarover de twee verzamelingen dus 'van mening verschillen'. De cruciale definitie luidt nu: theorie B lijkt meer op de ware theorie X dan theorie A als het symmetrisch verschil tussen B en X 'kleiner' is dan tussen A en X, in de zin dat de eerste verschilverzameling een echte deelverzameling vormt van de tweede.

2.3 Algemene evaluatie

Deze definitie stemt niet naar de letter maar wel naar de geest overeen met de definitie van Popper. Hij eiste dat theorie B meer ware en minder onware consequenties had dan theorie A. De structuralistische definitie kan op een analoge wijze geparafraseerd worden: theorie B laat meer (conceptuele) mogelijkheden terecht toe dan A en minder ten onrechte. Uitgaande van een eerste orde taal is het ook mogelijk aan te geven, zowel in termen van beweringen als mogelijkheden, waar het verschil precies zit en hoe het komt dat Poppers definitie wel en de structuralistische definitie geen schipbreuk lijdt (zie Kuipers:1982,1987b).

Het ruimte laten voor convergerende onware theorieën komt overigens niet door het centraal stellen van theoretische waarheid, waardoor X in de regel een verzameling met meer dan één (empirische) mogelijkheid is. Miller had al laten zien dat de (modeltheoretische variant van de) structuralistische definitie ook al voldoende ruimte schept als 'de waarheid' wordt opgevat als 'de ware beschrijving van de actuele wereld (mogelijkheid)' en dus als een verzameling met maar één mogelijkheid.⁵

Popper gebruikte in zijn definitie de parafraseringen 'dichter bij de waarheid' en 'meer gelijkend op de waarheid' enerzijds en 'beter corresponderend met de feiten' anderzijds door elkaar. Dit is opvallend omdat Popper elders niet alleen nadrukkelijk de correspondentietheorie van de waarheid onderschrijft, maar deze ook ziet als een 'interlevel theorie', een theorie die de niveaus van taal en werkelijkheid verbindt. De eerste twee parafraseringen zijn echter onmiskenbaar *intralevel* uitdrukkingen. Elders (Kuipers:1987b) heb ik beargumenteerd dat de intuïties die ten grondslag liggen aan de correspondentie-theorie van de waarheid niet alleen, zoals Popper in een soort *slip of the pen* al aangaf, als *intralevel* intuïties kunnen worden uitgelegd, maar dat daarbij het niveau van structuren nog geschikter is dan het niveau van beweringen. Aldus kan de structuralistische uitleg van waarheidsgelijkenis

5. Miller beschouwt mijn theoretische wending trouwens als een *definite improvement* (persoonlijke correspondentie).

gebruikt worden voor een demystificatie van de correspondentietheorie van de waarheid.

2.4 Het succes van theorieën

In deze uiteenzetting concentreer ik me echter op het verklaren van het succes van theorieën en kwesties die daar mee samenhangen. De structuralistische waarheidsgelijkenistheorie pretendeert een direct en ondubbelzinnig positief antwoord te kunnen geven op de uitdaging van Laudan. De volgende stelling kan namelijk bewezen worden: als een theorie meer lijkt op de waarheid dan een tweede dan zal de eerste theorie altijd succesvoller zijn dan de tweede (of tenminste even succesvol). Het gevolg is dat als een theorie succesvoller is dan een andere dit verklaard kan worden, met een beroep op genoemde stelling, door de hypothese dat de eerste theorie meer lijkt op de waarheid dan de tweede. Hierbij wordt niet alleen 'meer lijken op de waarheid', maar ook 'succesvoller' structuralistisch geïnterpreteerd. Dit laatste gebeurt op een wijze die overigens enerzijds goed aansluit bij denkbeelden over empirisch succes en empirische vooruitgang van Popper en Lakatos en anderzijds geheel past in de traditie van de logische empiristen, tot en met de hedendaagse vertegenwoordigers Van Fraassen (1980) en Laudan (1977).

Popper noemt als noodzakelijke (maar niet voldoende) voorwaarde voor vooruitgang dat de nieuwe theorie succesvoller moet zijn dan de voorganger. Hierbij wordt de nieuwe theorie succesvoller genoemd als de nieuwe theorie bepaalde experimentele feiten kan verklaren en wel: alle feiten die door de oude theorie verklaard werden, feiten die deze niet kon verklaren en feiten die zelfs leiden tot de falsificatie van de oude theorie. Lakatos heeft de strekking van deze eis als volgt opgenomen in zijn beroemde definitie van (theoretische en empirische) vooruitgang: de nieuwe theorie moet het succes van de oude theorie verklaren. Zoals bekend hebben Popper en Lakatos niet veel aandacht besteed aan de precieze aard van experimentele feiten. Soms lijkt het te gaan om experimentele algemene feiten, dus empirische wetten, soms staat de idee van tegenvoorbeelden centraal.

Soortgelijke onduidelijkheid is er in logisch empiristische hoek. Enerzijds wordt het succes van theorieën traditioneel uitgedrukt in termen van de positieve instanties, de voorbeelden, van de theorie. Anderzijds gaat het zeker ook om het verklaren van empirische wetten. Van Fraassen (1980) en Laudan (1977) zijn ook uitermate onduidelijk over wat nu precies 'empirisch adequate theorieën' respectievelijk 'empirische problemen' zijn.

Vanuit het structuralistisch perspectief dringt zich een technische

definitie van het succes van een theorie op die recht doet aan al deze kennelijk verstrengelde intuïties. Het succes van een theorie bestaat uit twee componenten die rechtstreeks samenhangen met het feit dat de empirische data twee aspecten hebben. Experimenteren betekent het realiseren van empirische mogelijkheden en dat leidt automatisch tot instanties die gerespecteerd moeten worden: het is goed als een theorie zo'n instantie toelaat, en dus als voorbeeld heeft, en fout als zij die niet toelaat, en dus als tegenvoorbeeld heeft.

Tegenover dit instantiële succes staat echter het verklarend succes. Een deel van de experimenten is in de regel expliciet bedoeld om al dan niet uit één van de theorieën afgeleide algemene hypothesen te toetsen en dit heeft geleid tot (voorlopige) aanvaarding van enkele van die hypothesen als empirische wetten: de te verklaren empirische wetten. Het is goed als een theorie zo'n wet verklaart en fout als dat niet zo is.

Beide succescomponenten kunnen in de structuralistische opzet precies gedefinieerd worden. Uit die definitie volgt onmiddellijk dat de twee componenten elkaar meestal tegenwerken: in de regel is een theorie die instantieel succesvoller is dan een andere verklarend juist minder succesvol, en omgekeerd. Dit stemt overeen met het intuïtieve inzicht dat een theorie waar alles inpast niets verklaart en dat een sterke theorie, die dus veel kan verklaren, een groter risico loop gefalsifieerd te worden dan een zwakke. Er is derhalve pas sprake van een echt succesvollere theorie als zij succesvoller is in één van de twee opzichten en tenminste succesbehoudend in het andere opzicht.

Welnu, als een theorie in deze gecombineerde zin succesvoller is dan een andere dan kan op basis van de hypothese dat de eerste theorie meer lijkt op de waarheid dan de tweede, bewezen worden dat de eerste in elk geval altijd in beide opzichten minstens even succesvol zal zijn dan de andere en dat de mogelijkheid van extra, nadrukkelijk bestaat. Met andere woorden, afgezien van de formele slag om de arm dat het extra succes nog niet te voorschijn hoeft te zijn gekomen, voldoet de explicatie van 'waarheidsgelijkenis' en 'succes' aan Laudans eis: grotere waarheidsgelijkenis impliceert meer succes.

2.5 Methodologische regels

De omgekeerde implicatie, meer succes impliceert grotere waarheidsgelijkenis, zou methodologie tot een eenvoudige zaak maken omdat meer succes in principe wel direct vaststelbaar is en grotere waarheidsgelijke-

nis niet. Maar het zal niet verbazen dat die implicatie niet geldt. Toch is het goed mogelijk methodologische regels van prescriptieve en heuristische aard te formuleren, die maximaal gebruikmaken van de wel geldige implicatie.

Het is niet moeilijk om te bewijzen dat de volgende *succesregel* (SR) 'Kies theorie B als theorie B succesvoller is dan theorie A' functioneel is voor het benaderen van de waarheid in de zwakke zin dat B niet minder op de waarheid kan lijken dan A. SR kan dus gerechtvaardigd worden als een prescriptieve regel en dient mijns inziens te worden opgevat als het keurmerk van wetenschappelijke rationaliteit.

De volgende regels betreffen heuristische regels waarvan gemakkelijk is na te gaan dat ze nieuwe toepassingen van SR stimuleren: de *inhoudsregel* 'Streef naar succesbehoudende versterking of, Popper ten spijt, succesbehoudende afzwakking van je theorie'⁶ en de *toetsingsregel* 'Streef naar de vaststelling van nieuwe tegenvoorbeelden van je theorie en van nieuwe wetten die niet verklaard worden door je theorie'. Tot slot, de *dialectische regel* voor twee theorieën die vanwege verdeeld succes aan SR ontsnappen: 'Streef naar een succesbehoudende synthese van twee aan SR ontsnappende theorieën'.

Het is belangrijk op te merken dat SR geen afleidingsregel is in de zin dat geconcludeerd zou worden dat de succesvollere theorie waar is (als hypothese, laat staan als theorie). Gieten we SR in de algemene vorm 'Kies de meest succesvolle theorie uit de beschikbare theorieën' dan kan hooguit de *tentatieve* conclusie getrokken worden dat de gekozen theorie het meeste lijkt op de waarheid, hetgeen de waarheidstoenadering-(WT-)inferentieregel genoemd zou kunnen worden ter onderscheiding van een waarheids-inferentieregel.

Waarheidsinferentie is geassocieerd met de zogenaamde *inference to the best explanation*, een regel die, volgens mijn interpretatie, voorschrijft of althans toestaat tentatief te concluderen, dat de meest succesvolle theorie waar is, aangenomen dat die theorie nog niet gefalsifieerd is. Verschillende bezwaren tegen en beperkingen van deze waarheidsinferentieregel zijn niet van toepassing op de voorgestelde WT-inferentieregel. In een discussie met Van Fraassen kwamen we samen tot de conclusie dat de belangrijkste verschillen de volgende zijn: ten eerste, de WT-regel is niet beperkt tot het geval dat de meest succesvolle theorie niet is gefalsifieerd, en ten tweede, het feit dat de notie 'het meeste lijken op de waarheid' altijd relatief is ten opzichte van andere (*in casu* de beschikbare) theorieën is vanzelfsprekend, maar dat

6. Waarbij het verzwakken respectievelijk versterken van een theorie neerkomt op het uitbreiden respectievelijk in krimpen van de verzameling conceptuele mogelijkheden.

‘waar-zijn’ relatief is ten opzichte van andere theorieën is vrijwel een *contradictio in terminis*.

Ter afsluiting van dit deel vat ik de hoofdlijnen van het structuralistisch convergentierealisme nog eens samen:

- een structuralistische benadering, waarin theorieën primair worden opgevat als verzamelingen structuren,
- gebaseerd op ‘theoretisch realisme’, volgens welke natuurwetenschappen niet primair gericht zijn op de karakterisering van de actuele wereld maar op karakterisering van de empirische mogelijkheden,
- met een explicatie van toenemende waarheidsgelijkenis in termen van een afnemend symmetrisch verschil met de ware theorie,
- met een tweezijdige explicatie van het succes van theorieën in termen van instantieel en verklarend succes,
- leidend tot een verklaring van toenemend succes als gevolg van grotere waarheidsgelijkenis,
- en tot methodologische regels die functioneel zijn voor het convergeren naar de waarheid.

3 Structuralistisch convergentierealisme: formeel

In dit deel zal ik eerst de formele kern van het structuralistisch convergentierealisme beknopt uiteenzetten. Veel uit het vorige deel zal daarbij in formeel-geëxpliceerde gedaante terugkeren. Vervolgens zal ik op grond daarvan een aanduiding geven van de extrapolatie naar gelaagde theorieën. Tot slot zal ik speculeren over de reikwijdte van de hele analyse. Bij dit deel wordt kennis van de elementaire verzamelingenleer voorondersteld.

3.1 Structuralistisch convergentierealisme als metawetenschappelijk programma

Het structuralistisch convergentierealisme (SCR) kan beschouwd worden als een (metawetenschappelijk explicatie-) onderzoeksprogramma gebaseerd op een idee (Kuipers:1989), kortweg een metaprogramma.

De *basisidee of -principe* (BP) van SCR is het volgende:

BP: uitgaande van een domein D van ‘natuurverschijnselen’ (toestanden, situaties, systemen) en een verzameling M van *conceptuele mogelijkheden*, ontworpen ter karakterisering van die verschijnselen, is er een unieke, tijds-onafhankelijke deelverzameling X van M, die de *empirische mogelijkheden* representeert; deze X vormt ‘de grote onbekende’ en dus het doel van het

theoriegerichte onderzoek in het domein.

De volgende opmerkingen kunnen BP verduidelijken.

- 1) M vormt het *conceptuele raamwerk* van een onderzoeksprogramma voor D. SCR is dus een metaprogramma over gewone onderzoeksprogramma's. M is, verzamelingen-theoretisch gesproken, een verzameling structuren van een bepaald similariteitstype.
- 2) Het is duidelijk dat X niet alleen afhankelijk is van M, maar tegelijkertijd ook van de (natuur-)werkelijkheid.
- 3) X representeert niet 'de geactualiseerde mogelijkheid' ('*the actual world*'), maar 'de actualiseerbare of empirische mogelijkheden', het gaat dus niet om de descriptieve waarheid, maar om de *theoretische waarheid*, het doel van theorievorming.
- 4) In overeenstemming met de beste tradities van empirische theorieën heeft de onderhavige metatheorie een theoretische term (X) die de centrale hypothetische entiteit van het metaprogramma vormt.
- 5) BP is in de gegeven formulering gebaseerd op deterministische vooronderstellingen (tengevolge van de dichotomie 'empirisch mogelijk/onmogelijk') en behoeft dus extrapolatie voor probabilistische contexten.
- 6) Extrapolatie naar sociale domeinen is problematisch omdat X, opgevat als een verzameling 'sociale mogelijkheden', niet tijdsafhankelijk zal zijn. De onder 5) bedoelde probabilistische extrapolatie zal daar geen soelaas voor bieden, omdat die wel tijdsafhankelijk is.

BP is de harde kern van een meta-onderzoeksprogramma (SCR) met vele verfijningen, specialisaties en toepassingen (zie onder andere Kuipers 1982, 1987 b/c/d, 1988). De verklaring van succesverschillen tussen theorieën is één zo'n toepassing. We zullen daar nu de weg voor vrijmaken door geschikte explicaties van enkele intuïtieve noties.

Een *theorie A* wordt opgevat als een deelverzameling A van M tezamen met de claim $A = X$, dat wil zeggen de claim dat A de (perfecte) karakterisering van de onbekende X is. Natuurlijk noemen we zo'n theorie waar als de claim waar is en onwaar als de claim onwaar is. Een direct gevolg hiervan is dat er precies één ware theorie is, namelijk theorie X. Om die reden mogen we X de (*theoretische*) *waarheid* noemen.

Theorie B *lijkt minstens evenveel op de waarheid als* theorie A dan en slechts dan als het symmetrisch verschil tussen de verzamelingen B en X ($B \Delta X =_{df} (B - X) \cup (X - B)$) een deelverzameling is van het symmetrisch verschil tussen A en X (dus $B \Delta X \subseteq A \Delta X$). Merk op dat het symmetrisch ver-

schil een soort afstand tussen de betreffende verzamelingen is. De gegeven definitie kan worden opgesplitst in twee clauses: de *instantiële clause* ($A \cap X \subseteq B \cap X$), die zegt dat alle mogelijkheden die terecht worden toegelaten door A ook worden toegelaten door B, en de *verklarende clause* ($B - X \subseteq A - X$), waarvan de benaming verderop zal worden gemotiveerd, die zegt dat alle mogelijkheden die ten onrechte worden toegelaten door B ook worden toegelaten door A.

De claim dat een theorie minstens evenveel lijkt op de waarheid dan een andere, in de boven gedefinieerde zin, is een perfect voorbeeld van een empirisch toetsbare, vergelijkende hypothese omdat zij, zoals we verderop zullen zien, voorspelt, en dus ook verklaart, dat de eerste altijd minstens zo succesvol zal zijn als de tweede. We zullen deze hypothese vanaf nu de *waarheidstoenadering (WT-) hypothese* noemen.

De definitie van 'minstens evenveel lijken op de waarheid' kan worden omgezet in een definitie van 'meer lijken op de waarheid' door te eisen dat het eerste symmetrisch verschil een *echte* deelverzameling moet zijn van het tweede. Hierdoor ontstaat de mogelijkheid van een serie onware theorieën die convergeert naar de waarheid.

3.2 De SCR-verklaring van succesverschillen tussen theorieën

Eerst moeten we een aantal voorbereidingen treffen voor het formuleren van succesbegrippen. Een *hypothese* H wordt opgevat als een zwakke variant van een theorie: H is een deelverzameling van M en de claim is dat X een deelverzameling is van H , $X \subseteq H$. Wederom geldt de afspraak dat hypothese H waar dan wel onwaar is als de claim waar respectievelijk onwaar is. Merk op dat er meer dan één ware hypothese is, terwijl er maar één ware theorie is en dat die neerkomt op 'de sterkste ware hypothese'.

De verzameling $H^*(A)$ van hypothesen die volgen uit theorie A kan worden gerepresenteerd als de verzameling van deelverzamelingen van M die A omvatten (formeel: $H^*(A) =_{df} \{H \subseteq M / A \subseteq H\}$). Het is nu niet moeilijk de 'brugstelling' te bewijzen (zie *Appendix 2*) dat de verklarende clause ($B - X \subseteq A - X$) uit de vorige paragraaf equivalent is met de conditie dat alle ware hypothesen die uit theorie A volgen ook uit theorie B volgen (formeel: $H^*(A) \cap H^*(X) \subseteq H^*(B) \cap H^*(X)$), hetgeen de naam van de clause rechtvaardigt.

De data waar onze theorieën rekenschap van moeten geven kunnen als volgt gerepresenteerd worden. Laat $R(t)$ de verzameling gerealiseerde mogelijkheden op tijdstip t aangeven, dat zijn dus ook de *te respecteren instanties*,

en laat $L^*(t)$ de verzameling geaccepteerde hypothesen op t aangeven, met andere woorden de *te verklaren wetten*. Op basis van $L^*(t)$ definiëren we de *sterkste wet* die verklaard moet worden als de hypothese $S(t)$ die geassocieerd is met de doorsnede van de verzamelingen die de geaccepteerde hypothesen constitueren.

Uiteraard is $L^*(t)$ en dus $S(t)$ op een of andere manier gebaseerd op $R(t)$. We mogen minimaal aannemen dat $R(t) \subseteq L$ voor iedere $L \in L^*(t)$, met het gevolg dat $R(t) \subseteq S(t)$. In wat volgt hebben we echter de veel sterkere *correcte data (CD-)hypothese* $R(t) \subseteq X \subseteq S(t)$ nodig, die zegt dat $R(t)$ geen fouten bevat en dat hypothese $S(t)$ waar is.

Het is nu mogelijk een explicatie te geven van het succes en de problemen van een theorie A op tijdstip t , waarbij de terminologie welbewust in de lijn van Laudans probleem-oplossingsmodel van wetenschap is gekozen. $A \cap R(t)$ geeft de verzameling gerealiseerde voorbeelden, het *instantieel succes* aan, terwijl $R(t) - A$ de verzameling gerealiseerde tegenvoorbeelden, de *instantiële problemen*, aangeeft. $H^*(A) \cap L^*(t)$ representeert de verklaarde wetten, het *verklarend succes* en $L^*(t) - H^*(A)$ de onverklaarde wetten, de *verklaringsproblemen*.

Voor vergelijkende oordelen over het succes van theorieën liggen nu de volgende twee clausules voor de hand. Theorie B is *instantieel minstens even succesvol* als theorie A dan en slechts dan als het instantieel succes van A een deelverzameling is van dat van B (formeel: $A \cap R(t) \subseteq B \cap R(t)$), hetgeen equivalent is aan de conditie dat de instantiële problemen van B een deelverzameling vormen van die van A . Theorie B is *verklarend minstens even succesvol* als theorie A dan en slechts dan als het verklarend succes van A een deelverzameling is van dat van B (formeel: $H^*(A) \cap L^*(t) \subseteq H^*(B) \cap L^*(t)$), hetgeen wederom equivalent is met de conditie dat de verklaringsproblemen van B een deelverzameling vormen van die van A . Op basis van een variant van de eerdergenoemde brugstelling is het gemakkelijk om te bewijzen dat de tweede (verklarend succes) clausule equivalent is aan de conditie dat de in het licht van $S(t)$ door B ten onrechte toegelaten mogelijkheden ook worden toegelaten door A (formeel: $B - S(t) \subseteq A - S(t)$).

De combinatie van de instantiële en de verklarende clausule vormt de definitie van de bewering dat de ene theorie op een bepaald moment minstens even succesvol is als een andere.

Nu zijn we bij de centrale redenering van het betoog. Als een theorie op een bepaald moment minstens even succesvol is als een andere dan kan dit worden afgeleid uit, en dus verklaard worden door, de volgende twee metahypothesen: de WT-hypothese dat de eerste minstens evenveel op de

waarheid lijkt als de tweede en de CD-hypothese dat de data correct zijn. Met andere woorden, de (centrale) stelling kan bewezen worden dat uit beide metahypothesen volgt dat de eerste theorie minstens even succesvol moet zijn als de tweede. Alle begrippen in deze stelling zijn uitgelegd en het bewijs (zie *Appendix 2*) is een kwestie van elementaire verzamelingen-theoretische manipulatie. Uit de stelling volgt ook meteen dat als beide metahypothesen waar zijn, de eerste theorie altijd minstens even succesvol zal blijven als de tweede, dus ook in het licht van nieuwe data.

In de regel zal de ene theorie niet alle instantieel succes van de ander omvatten en/of niet alle verklarend succes, laat staan beide vormen van succes. De idee ligt voor de hand dat de relatieve verdiensten dan verklaard kunnen worden op grond van een gedetailleerde analyse van de relatieve 'posities' ten opzichte van de waarheid, maar in dit geval is een algemene redenering niet mogelijk.

De centrale redenering maakt duidelijk dat en hoe empirisch vooruitgang mogelijk is binnen een conceptueel raamwerk M en een domein D . Het is van belang op te merken dat de specifieke WT-hypothese vooronderstelt dat BP waar is met betrekking tot $\langle D, M \rangle$. Deze toepassing van BP creëert als het ware de mogelijkheid dat er theorieën zijn die dichter bij de waarheid zijn dan andere en dat het feit dat theorieën succesvoller zijn dan andere daardoor veroorzaakt kan, maar niet hoeft, te worden. Met andere woorden, hoewel ieder specifiek voorbeeld van empirische vooruitgang verklaard wordt op basis van de corresponderende WT-hypothese, wordt het algemene verschijnsel van empirische vooruitgang in het kader van $\langle D, M \rangle$ verklaard door aan te nemen dat BP waar is met betrekking tot $\langle D, M \rangle$.

Twee opeenvolgende generalisaties brengen ons bij de verklaring van het succes van de natuurwetenschappen in het algemeen. Ten eerste de vooronderstelling dat BP waar is voor alle conceptuele raamwerken M met betrekking tot het natuurlijke domein D , of korter, dat BP waar is voor D . In de tweede plaats de vooronderstelling dat BP waar is voor alle natuurlijke domeinen. Natuurlijk claimen we niet dat deze generalisaties geen uitzonderingen hebben. Als ze in de meerderheid van de gevallen waar zijn dienen ze al hun doel.

3.3 Het succes van gelaagde theorieën

Tot nu toe lijkt het er misschien op dat ons conceptueel relatief vertrekpunt leidt tot een exotische vorm van relativistisch realisme, bijna een *contradictie in terminis*. Dit zou echter alleen het geval zijn als we verbanden tussen ver-

schillende conceptuele raamwerken voor hetzelfde domein zouden uitsluiten. In deze paragraaf zullen we de relatie behandelen tussen een *observationeel* en een *theoretisch* raamwerk voor eenzelfde domein. Hierbij wordt uiteraard aangenomen dat het onderscheid tussen observationele en theoretische componenten niet de klassieke, absolute vorm heeft, maar een genuanceerde, relatieve vorm.

Laat M_t de verzameling *conceptueel-theoretische mogelijkheden* aangeven en M_o de verzameling *conceptueel-observationele mogelijkheden*. We veronderstellen dat deze twee verzamelingen gerelateerd worden door een projectfunctie π van M_t op M_o , zodanig dat $\pi(x)$, technisch gesproken, een substructuur is van $x \in M_t$. Met andere woorden, elementen van M_t worden geacht ook de observationele componenten te bevatten. Als $A \subseteq M_t$ dan is πA per definitie de verzameling van alle projecties van elementen van A .

Toepassing van BP leidt direct tot het bestaan van unieke, tijdsafhankelijke deelverzamelingen $X(M_t) = X_t$ van M_t , genaamd de verzameling *empirisch-theoretische mogelijkheden*, en $X(M_o) = X_o$ van M_o , genaamd de verzameling *empirisch-observationele mogelijkheden*. Het is plausibel om bovendien het *principe van X-behoud* aan te nemen: $\pi(X_t) \subseteq X_o$. Dat $\pi(X_t)$ een deelverzameling moet zijn van X_o is alleen een semantische kwestie: als iets empirisch mogelijk is, dan moet het observeerbare deel ook empirisch mogelijk zijn.

In de onderhavige context van gelaagde theorieën wordt observationeel succes uiteraard uitgedrukt in termen van deelverzamelingen $R(t)$ en $S(t)$ van M_o die de verzameling *gerealiseerde empirisch-observationele mogelijkheden*, respectievelijk, *de sterkste ware wet* aangeven. De verklaring van het observationele succes moet in twee delen worden gesplitst.

Dat theorie B op observationeel niveau verklarend minstens even succesvol is als theorie A (formeel: $\pi B - S(t) \subseteq \pi A - S(t)$) volgt uit de hypothese dat 'B verklarend minstens evenveel lijkt op X_t als A' (formeel: $B - X_t \subseteq A - X_t$). Gebruikmakend van X-behoud impliceert deze hypothese immers direct dat πB verklarend minstens evenveel lijkt op X_o als πA (formeel: $\pi B - X_o \subseteq \pi A - X_o$), hetgeen op zijn beurt de bedoelde conclusie impliceert, aangenomen dat bovendien geldt dat alle te verklaren wetten waar zijn als hypothese ofwel, wat op hetzelfde neerkomt, dat $X_o \subseteq S(t)$.

Tengevolge van het veel-één karakter van projectie kan het corresponderende argument niet gebruikt worden voor de verklaring van het observationeel-instantieel succes. Aangenomen dat $R(t)$ deugt, dat wil zeggen $R(t) \subseteq X_o$, is de noodzakelijke en voldoende voorwaarde om te bewijzen dat theorie B, op observationeel niveau, instantieel minstens even succesvol is als theorie

A (formeel: $\pi A \cap R(t) \subseteq \pi B \cap R(t)$) de hypothese dat πB instantieel minstens evenveel lijkt op X_0 als πA (formeel: $\pi A \cap X_0 \subseteq \pi B \cap X_0$). Deze hypothese zegt in wezen dat de representationele kracht van B met betrekking tot X_0 minstens even groot is als die van A : voor alle $x \in X_0$ geldt als er een $a \in A$ is met $\pi(a) = x$ dan is er een $b \in B$ met $\pi(b) = x$. Uiteraard blijft meer instantieel succes op theoretisch niveau verklaarbaar in termen van instantieel grotere gelijkenis met de waarheid.

3.4 Speculaties over reikwijdte

Het principe van X -behoud is de eerste *constraint* die afstand schept tot een absoluut relativisme van conceptuele raamwerken. In deze laatste paragraaf zullen we enkele begrippen formuleren die het mogelijk maken om verdergaande non-relativistische posities te exploreren.

Een conceptueel raamwerk M voor domein D , met empirische mogelijkheden $X(M) \subseteq M$ op grond van BP, is een (*conceptueel*) *niveau* voor D als $X(M') = X(M)$ voor iedere M die M' omvat. De elementen van M' kunnen bijvoorbeeld een relatie als component bevatten waar de elementen van M een functie hebben die voldoende blijkt voor de karakterisering van X . Vanaf hier zijn M , M_0 en M_t conceptuele niveaus voor D , waarbij tussen M_0 en M_t , in de lijn van de vorige paragraaf, minimaal de volgende relatie bestaat:

M_t is een *superniveau* van M_0 (en M_0 een *subniveau* van M_t) als er een projectiefunctie π van M_t op M_0 is zodanig dat $\pi(x)$ een substructuur van $x \in M_t$ and $\pi X_t \subseteq X_0$, dat wil zeggen X -behoud, waarbij X_t en X_0 staan voor $X(M_t)$ respectievelijk $X(M_0)$.

Theorie A_t ($A_t \subseteq M_t$) *reproduceert* theorie A_0 ($A_0 \subseteq M_0$) als $\pi^{-1}A_0 =_{\text{df}} \{x \in M_t / \pi(x) \in A_0\} = A_t$, hetgeen, tengevolge van de algemeen geldigheid van $\pi^{-1}A_0 = A_0$, impliceert dat $\pi A_t = A_0$. Het is nu niet moeilijk om de volgende *communurabiliteitsstelling* te bewijzen: voor elk tweetal niveaus is er tenminste een gemeenschappelijk superniveau (eventueel op triviale wijze gedefinieerd door aaneenschakeling van componenten) waarop alle theorieën gereproduceerd en dus vergeleken kunnen worden.

M_t is *compleet* met betrekking tot M_0 als $\pi X_t \supseteq X_0$ (en dus, met X -behoud, $\pi X_t = X_0$). Merk op dat als M_t niet-refererende termen invoert deze waarschijnlijk incompleet is met betrekking tot M_0 . We zeggen dat M_t M_0 *reproduceert* als X_t X_0 reproduceert, dat wil zeggen $\pi^{-1}X_0 = X_t$. Hierbij is $\pi^{-1}X_0 \subseteq X_t$ de substantiële conditie bovenop X -behoud, want X -behoud ($\pi X_t \subseteq X_0$) impliceert direct $\pi^{-1}X_0 \subseteq X_t$.

Het is eenvoudig na te gaan dat als M_t compleet is met betrekking tot

M_0 dat M_t dan M_0 reproduceert. Merk ook op dat M_t M_0 slechts zal reproduceren als M_t alleen maar irrelevante componenten introduceert.

M_t is (*empirisch*) *rijker dan* M_0 als M_t compleet is met betrekking tot M_0 , maar de laatste niet slechts reproduceert. Formeel: $\pi X_t = X_0$ (implicierend dat $\pi^{-1}X_0 \supseteq X_t$), maar niet $\pi^{-1}X_0 \subseteq X_t$, en dus $\pi^{-1}X_0 = X_t$.

M is *niet-beperkbaar* als M rijker is dan al zijn subniveaus. M is *uitputtend* als M geen rijker superniveau heeft (en dus geen niet-beperkbaar superniveau). Merk op dat als M en M' beide uitputtend zijn (maar niet zodanig dat $M \subseteq / \supseteq M'$) dan zal de één een subniveau zijn van de ander. Tot slot, M is *optimaal* als M uitputtend en niet-beperkbaar is, ofwel, in suggestieve termen, als $X(M)$ de hele waarheid is en niets dan de waarheid.

In het gepresenteerde perspectief kan de klassieke *ideale taal assumptie* (ITA) als volgt geformuleerd worden: voor ieder domein bestaat een optimaal conceptueel niveau. Dit lijkt wel de sterkste niet-relativistische positie: het komt neer op extreem metafysisch realisme. ITA vooronderstelt natuurlijk BP en dus is ITA voor de sociale wetenschappen minstens zo problematisch als BP zelf. Maar ITA is zeker ook al problematisch voor de natuurwetenschappen. Hoewel ITA regelmatig vruchtbaar kan zijn als heuristisch principe wordt het volgende ermee conflicterende principe vaker verdedigd als richtlijn, bijvoorbeeld door Popper: hoewel er beperkbare niveaus bestaan zijn er geen uitputtende niveaus en dus geen optimale niveaus. Met andere woorden, voor ieder niveau is er een rijker niveau te vinden. Dit methodologisch principe kan het *verfijningsprincipe* genoemd worden. Iedere geslaagde toepassing ervan leidt niet alleen tot nieuwe typen van empirisch succes en dus van empirische vooruitgang, maar ook, in combinatie met het basisprincipe BP tot de verklaring van deze verschijnselen.⁷

4 Appendix 1: bewijs van de stelling van Miller en Tichý

Uitgangspunt is dat alle beschouwde theorieën eindig axiomatiseerbaar zijn en dus kunnen worden weergegeven door beweringen α , β , etc.. Met $C(\alpha)$ geven we de verzameling van alle consequenties van α weer. Zij θ de volledige en consistente waarheid, dat wil zeggen $C(\theta)$ bevat precies alle ware beweringen en iedere bewering α of zijn ontkenning $\neg\alpha$ behoort tot $C(\theta)$. In het eerste geval is α een ware bewering en geldt uiteraard dat $C(\alpha)$ een

7. Met dank aan dr. Henk Zandvoort en de redacteurs van dit boek voor het commentaar op de voorlopige versie.

deelverzameling van $C(\theta)$ is, in het tweede geval is α een onware bewering en geldt, alleen al vanwege α zelf, dat $C(\alpha) - C(\theta)$ niet leeg is. De verzameling onware beweringen vormt het complement van $C(\theta)$ en geven we aan met $F(\theta)$.

Poppers definitie, de stelling van Miller en Tichý en het bewijs komen nu op het volgende neer.

Definitie

β is dichterbij θ dan α dan en slechts dan als:

- (i) $C(\alpha) \cap C(\theta) \subseteq C(\beta) \cap C(\theta)$
- (ii) $C(\beta) \cap F(\theta) \subseteq C(\alpha) \cap F(\theta)$
- (iii) ' \subseteq ' tenminste éénmaal een echte deelverzameling.

Het is eenvoudig om na te gaan dat (i) zegt dat alle ware consequenties van α ook (ware) consequenties van β zijn en (ii) dat alle onware consequenties van β ook (onware) consequenties van α moeten zijn; (iii) voegt er aan toe dat β tenminste één extra ware consequentie moet hebben en/of tenminste één onware consequentie minder.

Stelling (Miller, Tichý, 1974):

Als β dichterbij θ is dan α dan is β waar, dat wil zeggen β is een element van $C(\theta)$, ofwel $C(\beta)$ is een deelverzameling van $C(\theta)$.

Bewijs:

Stel α , β en θ zijn zoals hierboven omschreven.

Uit (ii) volgt meteen $C(\beta) \cap F(\theta) \subseteq C(\alpha)$ en daaruit, met gebruikmaking van het feit dat $F(\theta)$ het complement is van $C(\theta)$, $C(\beta) \subseteq C(\alpha) \cup C(\theta)$. Daar $\beta \in C(\beta)$ volgt nu $\beta \in C(\alpha)$ en/of $\beta \in C(\theta)$. Stel nu dat β onwaar is. De tweede mogelijkheid is daarmee uitgesloten en er moet dus gelden $\beta \in C(\alpha)$.

Uit de volledigheid van θ volgt nu echter ook $\neg\beta \in C(\theta)$ en dus $\neg\theta \in C(\beta)$. Voorts volgt uit (i) meteen $C(\alpha) \cap C(\theta) \subseteq C(\beta)$ en, omdat de verzameling consequenties van een disjunctie gelijk is aan de doorsnede van de verzamelingen consequenties van de disjuncten afzonderlijk, volgt $C(\alpha \vee \theta) \subseteq C(\beta)$ en dus $\alpha \vee \theta \in C(\beta)$. Omdat al geldt $\neg\theta \in C(\beta)$ moet nu gelden $\alpha \in C(\beta)$.

Combineren we de twee tussenresultaten dan volgt dat α en β equivalent zijn. Hieruit volgt dat $C(\alpha)$ en $C(\beta)$ aan elkaar gelijk zijn, maar dat betekent dat niet voldaan kan zijn aan conditie (iii), want genoemde gelijkheid impliceert dat het ook in (i) en (ii) links en rechts om gelijke verzame-

lingen gaat. Conclusie: de (enige) onderweg gemaakte vooronderstelling, namelijk dat β onwaar is, kan niet juist zijn, met andere woorden β moet waar zijn. *Q.E.D.*

5 Appendix 2: bewijs van de brugstelling en de centrale stelling

Brugstelling: $H^*(A) \cap H^*(X) \subseteq H^*(B) \cap H^*(X) \Leftrightarrow B-X \subseteq A-X$

N.b. $H^*(A) =_{df} \{H \subseteq M / A \subseteq H\}$

Bewijs:

Eerst bewijzen we voor willekeurige A, B en C;

$$(i) \quad H^*(A) \cap H^*(B) = H^*(A \cup B)$$

$$(ii) \quad A \subseteq B \Leftrightarrow H^*(B) \subseteq H^*(A)$$

$$(iii) \quad A \subseteq B \cup C \Leftrightarrow A-B \subseteq C$$

(i) volgt uit de overweging dat een verzameling die zowel A als B omvat ook $A \cup B$ omvat en de overweging dat een verzameling die $A \cup B$ omvat zowel A als B omvat. (ii) volgt meteen uit de overweging dat als A een deelverzameling is van B dan omvatten alle verzamelingen die B omvatten ook A. (iii) tot slot is elementair.

Op grond van (i) geldt dat de linker bewering in de brugstelling equivalent is met $H^*(A \cup X) \subseteq H^*(B \cup X)$. Uit (ii) volgt dan meteen dat dit equivalent is met $B \cup X \subseteq A \cup X$. Dit is triviaal equivalent met $B \subseteq A \cup X$ en dat is volgens (iii) equivalent met $B-X \subseteq A$ en dat is triviaal equivalent met de rechter bewering in de brugstelling *Q.E.D.*

Centrale stelling:

Als (i) theorie B minstens evenveel lijkt op de waarheid als theorie A (WT-hypothese) en als (ii) de data correct zijn (CD-hypothese) dan (iii) is B minstens even succesvol als A.

Bewijs:

Volgens de definities in de tekst komt dit formeel neer op de redenering:

$$(i) \quad A \cap X \subseteq B \cap X \text{ en } B-X \subseteq A-X$$

$$(ii) \quad R(t) \subseteq X \text{ en } X \subseteq S(t)$$

$$-(iii) \quad A \cap R(t) \subseteq B \cap R(t) \text{ en } B-S(t) \subseteq A-S(t)$$

De 'kolommen' links en rechts kunnen onafhankelijk van elkaar bewezen worden. Bij wijze van voorbeeld nemen we de linkerkolom. Uit (ii)-links volgt direct $A \cap R(t) \subseteq A \cap X$. Dus volgt met (i)-links $A \cap R(t) \subseteq B \cap X$. Hieruit meteen $A \cap R(t) \subseteq B$ en omdat $A \cap R(t) \subseteq R(t)$ vanzelf geldt volgt (iii)-links. *Q.E.D.*

6 Beknopte gids tot de literatuur

Alle vier in Deel 1 aangeduide benaderingen van waarheidsgelijkenis en waarheidsconvergentie worden door de betreffende verdedigers uiteengezet in Kuipers (1987a). Daarnaast heeft Niiniluoto zijn aanpak zeer uitvoerig en systematisch gepresenteerd in een lijvig boek (1987). Ook Oddie heeft een boek gepubliceerd over de bijzondering van hem en Tichý (1986).

Voor een polemische beschouwing over het objectieve waarheidsbegrip in het verlengde van dit hoofdstuk zie mijn (1990).